

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/389,835A

DATE: 03/16/2001
TIME: 16:31:23

Input Set : A:\B0051084.txt
Output Set: N:\CRF3\03162001\I389835A.raw

ENTERED

6 <110> APPLICANT: Ruoho, Arnold E.
8 Geiser, Andrew H.
10 Krebs, Mark
12 Sievert, Mike
16 <120> TITLE OF INVENTION: BACTERIORHODOPSIN/G PROTEIN-COUPLED RECEPTOR CHIMERAS
21 <130> FILE REFERENCE: 96429/9079
25 <140> CURRENT APPLICATION NUMBER: US/09/389,835A
27 <141> CURRENT FILING DATE: 1999-09-03
31 <150> PRIOR APPLICATION NUMBER: 60/098,950
33 <151> PRIOR FILING DATE: 1998-09-03
37 <160> NUMBER OF SEQ ID NOS: 53
41 <170> SOFTWARE: Word 97 (DOS text file)
45 <210> SEQ ID NO: 1
47 <211> LENGTH: 1626
49 <212> TYPE: DNA
51 <213> ORGANISM: Halobacterium salinarium
55 <220> FEATURE:
57 <221> NAME/KEY: CDS
59 <222> LOCATION: (394)..(1182)
63 <400> SEQUENCE: 1
65 ggatccgacg tgaagatggg gctcccgatg ggtgcaaccg tgaagtcctg cacggctgcg 60
69 tcacgacagg agccgaccag cgacacccag aaggtgcgaa cggttgagtg ccgcaacgat 120
73 cacgagtttt tcgtgcgctt cgagtggtaa cagcgtgca cgcctcgact tcaccgcggg 180
77 tgtttcgacg ccagccggcc gttgaaccag caggcagcgg gcatttcaca gccgctgtgg 240
81 ccacacact cgggtggggtg cgctattttg gtatggtttg gaatccgcgt gtcggctccg 300
85 tgtctgacgg ttcctcggtc taaattccgt cagcagcgta ccatactgat tgggtcgtag 360
89 agttacacac atatcctcgt taggtactgt tgc atg ttg gag tta ttg cca aca 414
91 Met Leu Glu Leu Leu Pro Thr
93 1 5
97 gca gtg gag ggg gta tcg cag gcc cag atc acc gga cgt ccg gag tgg 462
99 Ala Val Glu Gly Val Ser Gln Ala Gln Ile Thr Gly Arg Pro Glu Trp
101 10 15 20
105 atc tgg cta gcg ctc ggt acg gcg cta atg gga ctc ggg acg ctc tat 510
107 Ile Trp Leu Ala Leu Gly Thr Ala Leu Met Gly Leu Gly Thr Leu Tyr
109 25 30 35
113 ttc ctc gtg aaa ggg atg ggc gtc tcg gac cca gat gca aag aaa ttc 558
115 Phe Leu Val Lys Gly Met Gly Val Ser Asp Pro Asp Ala Lys Lys Phe
117 40 45 50 55
121 tac gcc atc acg acg ctc gtc cca gcc atc gcg ttc acg atg tac ctc 606
123 Tyr Ala Ile Thr Thr Leu Val Pro Ala Ile Ala Phe Thr Met Tyr Leu
125 60 65 70
129 tcg atg ctg ctg ggg tat ggc ctc aca atg gta ccg ttc ggt ggg gag 654
131 Ser Met Leu Leu Gly Tyr Gly Leu Thr Met Val Pro Phe Gly Gly Glu
133 75 80 85
137 cag aac ccc atc tac tgg gcg cgg tac gct gac tgg ctg ttc acc acg 702
139 Gln Asn Pro Ile Tyr Trp Ala Arg Tyr Ala Asp Trp Leu Phe Thr Thr
141 90 95 100

RECEIVED

MAR 23 2001

TECH CENTER 1600/2000

RAW SEQUENCE LISTING

DATE: 03/16/2001

PATENT APPLICATION: US/09/389,835A

TIME: 16:31:23

Input Set : A:\B0051084.txt

Output Set: N:\CRF3\03162001\I389835A.raw

```

145 ccg ctg ttg ttg tta gac ctc gcg ttg ctc gtt gac gcg gat cag gga 750
147 Pro Leu Leu Leu Leu Asp Leu Ala Leu Leu Val Asp Ala Asp Gln Gly
149 105 110 115
153 acg atc ctt gcg ctc gtc ggt gcc gac ggc atc atg atc ggg acc ggc 798
155 Thr Ile Leu Ala Leu Val Gly Ala Asp Gly Ile Met Ile Gly Thr Gly
157 120 125 130 135
161 ctg gtc ggc gca ctg acg aag gtc tac tcg tac cgc ttc gtg tgg tgg 846
163 Leu Val Gly Ala Leu Thr Lys Val Tyr Ser Tyr Arg Phe Val Trp Trp
165 140 145 150
169 gcg atc agc acc gca gcg atg ctg tac atc ctg tac gtg ctg ttc ttc 894
171 Ala Ile Ser Thr Ala Ala Met Leu Tyr Ile Leu Tyr Val Leu Phe Phe
173 155 160 165
177 ggg ttc acc tcg aag gcc gaa agc atg cgc ccc gag gtc gca tcc acg 942
179 Gly Phe Thr Ser Lys Ala Glu Ser Met Arg Pro Glu Val Ala Ser Thr
181 170 175 180
185 ttc aaa gta ctg cgt aac gtt acc gtt gtg ttg tgg tcc gcg tat ccc 990
187 Phe Lys Val Leu Arg Asn Val Thr Val Val Leu Trp Ser Ala Tyr Pro
189 185 190 195
193 gtc gtg tgg ctg atc ggc agc gaa ggt gcg gga atc gtg ccg ctg aac 1038
195 Val Val Trp Leu Ile Gly Ser Glu Gly Ala Gly Ile Val Pro Leu Asn
197 200 205 210 215
201 atc gag acg ctg ctg ttc atg gtg ctt gac gtg agc gcg aag gtc ggc 1086
203 Ile Glu Thr Leu Leu Phe Met Val Leu Asp Val Ser Ala Lys Val Gly
205 220 225 230
209 ttc ggg ctc atc ctc ctg cgc agt cgt gcg atc ttc ggc gaa gcc gaa 1134
211 Phe Gly Leu Ile Leu Leu Arg Ser Arg Ala Ile Phe Gly Glu Ala Glu
213 235 240 245
217 gcg ccg gag ccg tcc gcc ggc gac ggc gcg gcc gcg acc agc gac tga 1182
219 Ala Pro Glu Pro Ser Ala Gly Asp Gly Ala Ala Ala Thr Ser Asp
221 250 255 260
225 tcgcacacgc aggacagccc cacaaccggc gcggcttttc aacgacacac gatgagtcgc 1242
229 ccaactcggtc ttgtactcgc acgatcgcg gcgacggcg acgccgacgg cgactttcca 1302
233 gcgtcgctca acaggctggc tgtcgtcgcg ctcgctggtg cggctctcgt cgggtcgggcg 1362
237 ggtctgttcg ccgtgccgtt cctgcggtcg ttcggcatga cgttttgga agcgttcacc 1422
241 gttgttggtg tctccgagtt cgtctcggcc atcgtggcgg ccctcgcggg ctaccacctc 1482
245 tacaccacgc ccgacgacta gcagggcccg ctggcgagcc atcactcccg ctgtggcgag 1542
249 gcgacggccg ttctgtaccg ctaccgccg ccgggagtcg gggacatcg cggggcgatg 1602
253 cgcacgaac ggtcaccgg atcc 1626
259 <210> SEQ ID NO: 2
261 <211> LENGTH: 262
263 <212> TYPE: PRT
265 <213> ORGANISM: Halobacterium salinarium
269 <400> SEQUENCE: 2
271 Met Leu Glu Leu Leu Pro Thr Ala Val Glu Gly Val Ser Gln Ala Gln
273 1 5 10 15
277 Ile Thr Gly Arg Pro Glu Trp Ile Trp Leu Ala Leu Gly Thr Ala Leu
279 20 25 30
283 Met Gly Leu Gly Thr Leu Tyr Phe Leu Val Lys Gly Met Gly Val Ser
285 35 40 45

```

RAW SEQUENCE LISTING

DATE: 03/16/2001

PATENT APPLICATION: US/09/389,835A

TIME: 16:31:23

Input Set : A:\B0051084.txt

Output Set: N:\CRF3\03162001\I389835A.raw

```

289 Asp Pro Asp Ala Lys Lys Phe Tyr Ala Ile Thr Thr Leu Val Pro Ala
291      50      55      60
295 Ile Ala Phe Thr Met Tyr Leu Ser Met Leu Leu Gly Tyr Gly Leu Thr
297 65      70      75      80
301 Met Val Pro Phe Gly Gly Glu Gln Asn Pro Ile Tyr Trp Ala Arg Tyr
303      85      90      95
307 Ala Asp Trp Leu Phe Thr Thr Pro Leu Leu Leu Leu Asp Leu Ala Leu
309      100      105      110
313 Leu Val Asp Ala Asp Gln Gly Thr Ile Leu Ala Leu Val Gly Ala Asp
315      115      120      125
319 Gly Ile Met Ile Gly Thr Gly Leu Val Gly Ala Leu Thr Lys Val Tyr
321      130      135      140
325 Ser Tyr Arg Phe Val Trp Trp Ala Ile Ser Thr Ala Ala Met Leu Tyr
327 145      150      155      160
331 Ile Leu Tyr Val Leu Phe Phe Gly Phe Thr Ser Lys Ala Glu Ser Met
333      165      170      175
337 Arg Pro Glu Val Ala Ser Thr Phe Lys Val Leu Arg Asn Val Thr Val
339      180      185      190
343 Val Leu Trp Ser Ala Tyr Pro Val Val Trp Leu Ile Gly Ser Glu Gly
345      195      200      205
349 Ala Gly Ile Val Pro Leu Asn Ile Glu Thr Leu Leu Phe Met Val Leu
351      210      215      220
355 Asp Val Ser Ala Lys Val Gly Phe Gly Leu Ile Leu Leu Arg Ser Arg
357 225      230      235      240
361 Ala Ile Phe Gly Glu Ala Glu Ala Pro Glu Pro Ser Ala Gly Asp Gly
363      245      250      255
367 Ala Ala Ala Thr Ser Asp
369      260
375 <210> SEQ ID NO: 3
377 <211> LENGTH: 20
379 <212> TYPE: DNA
381 <213> ORGANISM: Artificial Sequence
385 <220> FEATURE:
387 <223> OTHER INFORMATION: Description of Artificial Sequence:oligonucleotide
389      primer
393 <400> SEQUENCE: 3
395 cgcgatatcca gtcgtgtggc 20
401 <210> SEQ ID NO: 4
403 <211> LENGTH: 20
405 <212> TYPE: DNA
407 <213> ORGANISM: Artificial Sequence
411 <220> FEATURE:
413 <223> OTHER INFORMATION: Description of Artificial Sequence:oligonucleotide
415      primer
419 <400> SEQUENCE: 4
421 cctcctgagg agtcgtgcga 20
427 <210> SEQ ID NO: 5
429 <211> LENGTH: 91
431 <212> TYPE: DNA

```

RAW SEQUENCE LISTING DATE: 03/16/2001
 PATENT APPLICATION: US/09/389,835A TIME: 16:31:23

Input Set : A:\B0051084.txt
 Output Set: N:\CRF3\03162001\I389835A.raw

```

433 <213> ORGANISM: Artificial Sequence
437 <220> FEATURE:
439 <223> OTHER INFORMATION: Description of Artificial Sequence:oligonucleotide
441     primer
445 <400> SEQUENCE: 5
447 atcctgtacg tgctgttctt cgggttcacc gtcaaggagg cggcggcgca gcagcaggag 60
451 tcggcgacga cgcagaaggc ggagaaggag g                               91
457 <210> SEQ ID NO: 6
459 <211> LENGTH: 96
461 <212> TYPE: DNA
463 <213> ORGANISM: Artificial Sequence
467 <220> FEATURE:
469 <223> OTHER INFORMATION: Description of Artificial Sequence:oligonucleotide
471     primer
475 <400> SEQUENCE: 6
477 cgggatacgc ggaccacaac acaacggtaa cgttacgcag tactttgaac gtggatgcga 60
481 cctccatgcg cgtgacctcc ttctccgcct tctgcg                               96
487 <210> SEQ ID NO: 7
489 <211> LENGTH: 26
491 <212> TYPE: DNA
493 <213> ORGANISM: Artificial Sequence
497 <220> FEATURE:
499 <223> OTHER INFORMATION: Description of Artificial Sequence:oligonucleotide
501     primer
505 <400> SEQUENCE: 7
507 gtacatcctg tacgtgctgt tcttcg                               26
513 <210> SEQ ID NO: 8
515 <211> LENGTH: 19
517 <212> TYPE: DNA
519 <213> ORGANISM: Artificial Sequence
523 <220> FEATURE:
525 <223> OTHER INFORMATION: Description of Artificial Sequence:oligonucleotide
527     primer
531 <400> SEQUENCE: 8
533 acgacgggat acgcggacc                               19
539 <210> SEQ ID NO: 9
541 <211> LENGTH: 22
543 <212> TYPE: DNA
545 <213> ORGANISM: Artificial Sequence
549 <220> FEATURE:
551 <223> OTHER INFORMATION: Description of Artificial Sequence:oligonucleotide
553     primer
557 <400> SEQUENCE: 9
559 atcctgtacg tgctgttctt cg                               22
565 <210> SEQ ID NO: 10
567 <211> LENGTH: 15
569 <212> TYPE: DNA
571 <213> ORGANISM: Artificial Sequence
575 <220> FEATURE:

```

RAW SEQUENCE LISTING DATE: 03/16/2001
 PATENT APPLICATION: US/09/389,835A TIME: 16:31:23

Input Set : A:\B0051084.txt
 Output Set: N:\CRF3\03162001\I389835A.raw

```

577 <223> OTHER INFORMATION: Description of Artificial Sequence:oligonucleotide
579     primer
583 <400> SEQUENCE: 10
585 cgggatacgc ggacc                               15
591 <210> SEQ ID NO: 11
593 <211> LENGTH: 83
595 <212> TYPE: DNA
597 <213> ORGANISM: Artificial Sequence
601 <220> FEATURE:
603 <223> OTHER INFORMATION: Description of Artificial Sequence:oligonucleotide
605     primer
609 <400> SEQUENCE: 11
611 atcctgtacg tgctgttctt cgggttcacc gcgcgctccc acacgcgcaa gatctccacg 60
615 ctcccgcgcg cgaacatgaa ggg                               83
621 <210> SEQ ID NO: 12
623 <211> LENGTH: 75
625 <212> TYPE: DNA
627 <213> ORGANISM: Artificial Sequence
631 <220> FEATURE:
633 <223> OTHER INFORMATION: Description of Artificial Sequence:oligonucleotide
635     primer
639 <400> SEQUENCE: 12
641 cgggatacgc ggaccacaac acaacggtaa cgttacgcag tactttgaac gtggatgcga 60
645 cgcccttcat gttcg                               75
651 <210> SEQ ID NO: 13
653 <211> LENGTH: 89
655 <212> TYPE: DNA
657 <213> ORGANISM: Artificial Sequence
661 <220> FEATURE:
663 <223> OTHER INFORMATION: Description of Artificial Sequence:oligonucleotide
665     primer
669 <400> SEQUENCE: 13
671 gggttcaccc aggtcttcta cctcatccgc aagcagctga caagaaggtc tccgcgtcct 60
675 ccggcgaccc gcagaagtac tacggcaag                               89
681 <210> SEQ ID NO: 14
683 <211> LENGTH: 90
685 <212> TYPE: DNA
687 <213> ORGANISM: Artificial Sequence
691 <220> FEATURE:
693 <223> OTHER INFORMATION: Description of Artificial Sequence:oligonucleotide
695     primer
699 <400> SEQUENCE: 14
701 cacaacggta acgttacgca gtactttgaa cgtggatgcg acggacttcg cgatcttgag 60
705 ctcccttgccg tagtacttct gcgggtcgcc                               90
711 <210> SEQ ID NO: 15
713 <211> LENGTH: 84
715 <212> TYPE: DNA
717 <213> ORGANISM: Artificial Sequence
721 <220> FEATURE:

```

VERIFICATION SUMMARY

DATE: 03/16/2001

PATENT APPLICATION: US/09/389,835A

TIME: 16:31:24

Input Set : A:\B0051084.txt

Output Set: N:\CRF3\03162001\I389835A.raw